Chromatography: A Quantified Visualization of Movies

Kyung yul Kevin Lim
University of California San Diego
La Jolla, California
kyl063@eng.ucsd.edu

Hyeonsu Kang
University of California San Diego
La Jolla, California
hyk149@eng.ucsd.edu

Jennifer He
University of California San Diego
La Jolla, California
p3he@eng.ucsd.edu

ABSTRACT
Film production and consumption has been on an increasing trend and will not slow down in the near future. Existing methods for studying film and media is not prepared to scale to this trend, and in fact has already fallen behind. This paper explores a way of studying film through a quantified way and use information visualization so that the film maker or critic can compare and contrast multiple films at the same time. We gather data about the film in three main ways: extracting color palette, parsing the script, and visualizing the audio. This visualized data is then used to make macro trends of film making clearer and also start a discussion of a field of quantified film studies.

Author Keywords
information visualization; film studies; media studies; color palette; film script; film audio

INTRODUCTION
Film has grown to become a large part of entertainment culture and economy, and doesn't seem to be slowing down at all. PricewaterhouseCoopers projects that by 2019, the film entertainment industry's revenue will have grown to 35.33 billion USD from 29.15 billion USD in 2015 [12]. Despite such prevalence of statistics about the economics of the film industry, there is a distinct lack of quantified information and analysis of the content of film itself. This is a huge oversight in the part of film studies and leaves a gap for the filmmaker, the film critic or film historian to learn more about the craft. No single person can process and retain all the information in multiple—or even a single—films at once due to the density of data in a film and analyzing meaningful comparisons is that much harder.

We implore that media studies needs to move towards a quantified paradigm to achieve a deeper level of research and analysis of films. Our project Chromatography sheds light on this underexplored area and attempts to create a visual means of exploring such quantified information. We've chosen three aspects of data extracted from films to explore in this paper: the color palette, character networks, and audio data.

PREVIOUS WORKS
We are not the first group of researchers to ask for a more quantitative approach to film studies. Nick Redfern has implored that the only sensible way to understand cinema is to apply a more quantitative approach that utilizes computational media aesthetics [14]. In many ways the spirit of our research follows from his call to action for film studies to become relevant to the wider world.

Cutting et. al had taken a quantitative approach to analyzing 150 films over 75 years and found that the language of film has evolved to better hold the audienceâĂŹs attention by becoming quicker in shot length and visual activity and darker in luminance [7]. Cutting et. al's methods of abstraction of film is inspiring, but their work doesn't provide an interactive platform for users to discover new insight about the films, besides the ones provided in the paper. In this era of interactive information visualizations, we believe this to be a missed opportunity and aim to address this by providing an interactive platform with the film data.

In terms of pre-existing interactive visualizations of film, there are many examples on the web such as Movie Barcode, The Colors of Motion, and Cinemetrics [2, 5, 3]. Even though they are aesthetically pleasing, we argue they do not give much insight into the actual narrative structure of the films being inspected. The issue is that they all work with average color of a frame and this is too coarse grain of an information for the viewer to recall the film from.

A more informative level of analysis is done by Yeh et al. through extracting a "tempo" from audio and visual cues of the movie [17]. Yeh et al.'s works make significant events from the movies more obvious than the previously cited 3 visualizations. However, their visualizations are limited to showing a single film, when there can be much value in comparing of many films at once. Our attempt at visualization addresses exploring multiple movies at once, much like Cutting had done.

The lack of multi-movie comparison seems to be a common point of criticism we've found with many of the previous works we've discovered. Even David Robinson's interesting visualization of character networks is limited to a single movie and is an ad hoc attempt that does not generalize to other movies, or compares across movies [15]. However, we did find one particular example from Kevin Ferguson's analysis of 54 Disney films that compared all those movies' hue and brightness well, and took inspiration from it [8].

To summarize, many pre-existing attempts to quantify movie information is lacking in at least one aspect of scalability, narrative significance, or interactivity. The main goal of our methodology aimed to solve these limitations.

METHODOLOGY

In order to have our work scalable, first we needed a pipeline for gathering many movie data. We ripped DVDs that we owned or could borrow using the freely available programs DVDFab 9 and VLC Media Player. This movie file would then be crawled by a python script for a sample of 200 frames (about 30s frequency in a movie with 100 minute running time), subtitle, and audio data. External script databases were also used to augment the narrative information [1].

The three main data we will be visualizing are color palette information, script and subtitle data, and audio volume.

After these data are gathered, transformed and visualized an in-browser javascript application displays the visualization, that will allow a user to dynamically isolate colors, or characters, and even compare two movies side by side. The types of interactions will focus on features that allow comparison of the data we have gathered so that users can gain insight on subjects such as director styles, similar or different color usage across movie genres or periods.

Color Palette

Color in film is a particularly important narrative component, as much as the characters and setting. The language of film uses color in association with characters to indicate a change in their arc, or to foreshadow an event. Since master filmmakers have used such devices over a long period of time, we believe there is a pattern and structure that we can make apparent through visualization.

In order to connect color with narrative intent, it is crucial to track color as a character of its own. Previous works such as Movie Barcode [2], The Colors of Motion [5], or Cinemetrics [3] fail to capture this information by naively averaging all the pixel values together. Instead, our method is to track individual colors throughout the movie by explicitly finding the palette of each of our analyzed frames, and keeping those colors separate as to not lose the fine grain information of individual colors.

The class of algorithms for extracting a palette from an image—referred to as Color Quantization—has existed in the computer graphics literature for a while and originally invented to display images on a computer when graphics memory used to be limited. Though the median cut clustering algorithm is the most popular method [9], we chose to modify a pre-existing python implementation of k-means clustering for our palette extraction [11]. Since we do not have the processing power to extract every frame's color, we only sample 200 frames from each film—as mentioned above.

To conclude this section, our hypothesis is that there are meaningful ways that color is used in the language of storytelling in film, and we can visualize these better through the use of tracking individual colors as opposed to average or median colors.

Script Data

High Complexity in Analyzing Script Data

Designing an exploratory visualization tool of rich movie data is an exploration in itself. The structural complexity involved in designing such a visualization has 3 dimensions: (1) the complexity and large quantity of a single movie data (e.g. high-resolution visual and audio information and high-volume subtitle data), (2) spanning irregularities across different movies and (3) the diversity in arranging different views of visualization as well as in designing different modes of interactivity across the views and between a user and the views.

Particularly, analyzing subtitle data is challenging due to its varying formats and the large quantity of text. A typical script analysis in line with its corresponding visual sequence requires scene detection and segmentation, image/text alignment, and screenplay/closed captions alignment [6] (Figure 1). Different scene detection and segmentation mechanisms [10], [4], [13] can have varying granularity depending on the purpose of analysis. For sub-story discovery or movie summarization, for instance, detecting semantically continuous scene boundaries is important [16]. The image/text alignment stage involves matching each shot of the movie to the corresponding text segment in the screenplay, and finally, the screenplay/closed captions alignment stage maps the screenplay text segments to the corresponding parts in subtitles.

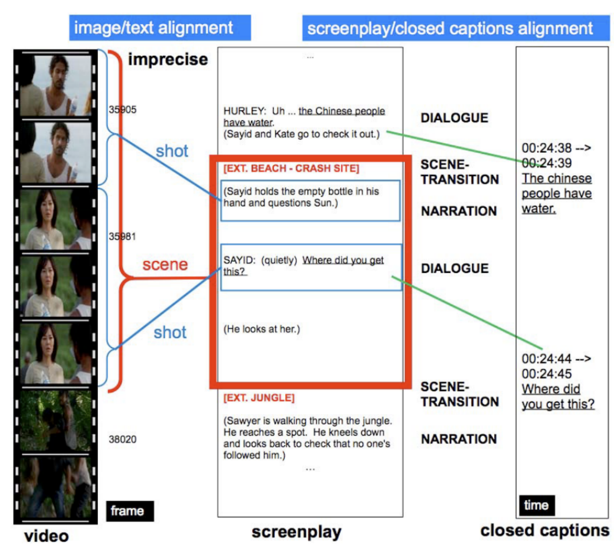


Figure 1: Video-text alignment, Cour et al.

The more complicated the method of processing script data is, the closer the analysis could bring us to a specific use of it. However, in analyzing script data in conjunction with other modes of audio-visual input, endlessly refining the precision of the analysis could be problematic, as other inputs' analytic precision should also be aligned in parallel. In Chromatography, we have adopted a rather simple but uniformly applicable speaker-content-scene model in analyzing subtitle data. The description of this process is as follows (Algorithm 1).

Algorithm 1 Subtitle Processing

S: The set of subtitle text segments
IF: The set of image frames
T: The set of uniform sampling time stamps $\{t_k\}$ of audio-visual input
procedure SUBTITLE PROCESSING(S, IF, T)
  for each $t_k \in T$ **do**
    - Clean S_{t_k}'s by removing non-character-communication descriptions
    - Regularize S_{t_k}'s by aggregating content to corresponding characters and by annotating character names and time stamps.
    - Map S_{t_k}'s to the corresponding group of IF_{t_k}'s using the encoded time stamps
  end for
  - Extract unique characters in each IF_{t_k}
  - Aggregate characters across different IF_{t_k}'s and map characters to speaker names to form a character-speaker-scene matrix
end procedure

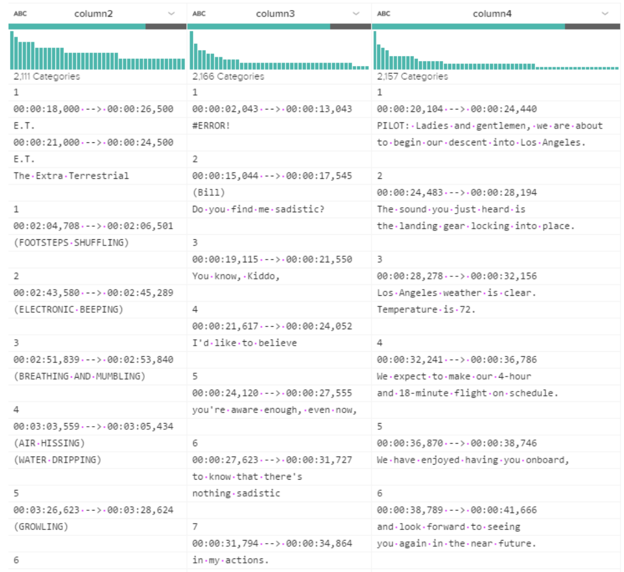
The main differences between our algorithm and the general subtitle processing procedure described earlier are (1) the simplification of the scene boundary detection algorithm and (2) the removal of the screenplay-closed caption alignment process. Instead, by uniformly sampling movie frames, we greatly simplified the notion of and computation involved in scene boundaries. At the same time, we may not have compromised a comparative level of accuracy by sampling 200 times, which gives around a 27-second of scene duration for movies of an-hour-and-a-half running time. The research of the effectiveness in extracting semantic scene boundaries to exposing the narrative structure of movies is a research topic for the future.

Cleaning and Regularization of Data

In Chromatography, the data cleaning and regularization process was necessary for aggregating characters' co-appearances within different scenes. Due to the variability in subtitle data within and across movies, this process required special care. For instance, in Figure 2, different formats of denotations in meta-data (e.g. descriptions of audio/visual effects) that appear across different movies can end up necessitating individual modifications in pre-processing scripts, and the problem looms even larger when the format is inconsistent within the movies. In addition to that, embedded erroneous data can reduce the degree of available automation in processing.

Visualizing Character Co-appearances

After cleaning and regularizing the script data, we can visualize the character-speaker-scene matrix in different ways. We present three alternatives (a dendrogram (Figure 3), a co-appearance heatmap (Figure 4) and a line-plot (Figure 5)) below.



column2	column3	column4
2,111 Categories	2,166 Categories	2,157 Categories
1	1	1
00:00:18,000 --> 00:00:26,500	00:00:02,043 --> 00:00:13,043	00:00:20,104 --> 00:00:24,440
E.T.	#ERROR!	PILOT: Ladies and gentlemen, we are about
00:00:21,000 --> 00:00:24,500		to begin our descent into Los Angeles.
E.T.	2	
The Extra Terrestrial	00:00:15,044 --> 00:00:17,545	2
	(Bill)	00:00:24,483 --> 00:00:28,194
1	Do you find me sadistic?	The sound you just heard is
00:02:04,708 --> 00:02:06,501		the landing gear locking into place.
(FOOTSTEPS SHUFFLING)	3	
	00:00:19,115 --> 00:00:21,550	3
2	You know, Kiddo,	00:00:28,278 --> 00:00:32,156
00:02:43,580 --> 00:02:45,289		Los Angeles weather is clear.
(ELECTRONIC BEEPING)	4	Temperature is 72.
	00:00:21,617 --> 00:00:24,052	
3	I'd like to believe	4
00:02:51,839 --> 00:02:53,840		00:00:32,241 --> 00:00:36,786
(BREATHING AND MUMBLING)	5	We expect to make our 4-hour
	00:00:24,120 --> 00:00:27,555	and 18-minute flight on schedule.
4	you're aware enough, even now,	
00:03:03,559 --> 00:03:05,434		5
(AIR HISSING)	6	00:00:36,870 --> 00:00:38,746
(WATER DRIPPING)	00:00:27,623 --> 00:00:31,727	We have enjoyed having you onboard,
	to know that there's	
5	nothing sadistic	6
00:03:26,623 --> 00:03:28,624		00:00:38,789 --> 00:00:41,666
(GROWLING)	7	and look forward to seeing
	00:00:31,794 --> 00:00:34,864	you again in the near future.
6	in my actions.	

Figure 2: Different Formats of Subtitles

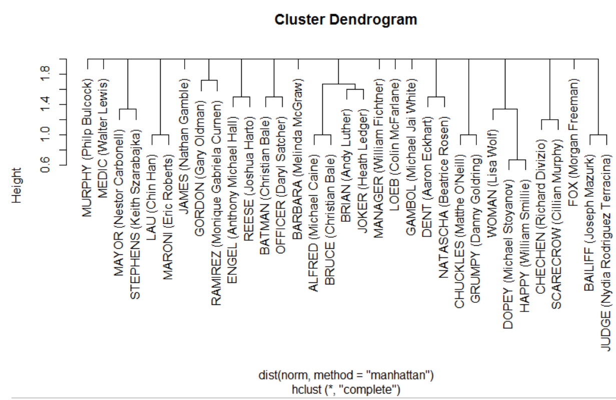


Figure 3: A Dendrogram of Characters in *The Dark Knight*

All of the three ways serve the basic goal of visualizing character co-appearances, but their focus are on different aspects. For instance, as shown in color variations in the co-appearance heatmap and in the varying heights of the dendrogram, the visualization can effectively represent the degree of co-appearance, or as the dot-plot of characters suggests, it can spread out the development of co-appearances over time, effectively establishing a continuum of time within the visualization of data. On the other hand, in rendering dynamic interactivity of the visualization, using the line-plot approach can have a benefit over the other two as it requires a minimal change in the visual – specifically, by adding a vertical bar on the graph – to indicate the flow of time.

However, since the line-plot spreads the co-appearance patterns over time, it can be useful to have a cumulative network model of characters to grasp the complexity of character relations developed toward the end of the movie. Therefore, we have adopted a slider control to let users change the scenes and a (cumulative) network visualization on top of the line

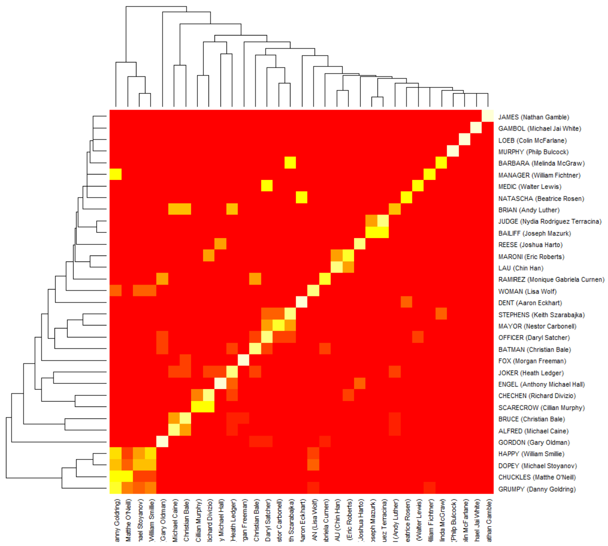


Figure 4: A Co-appearance Heatmap of Characters in *The Dark Knight*

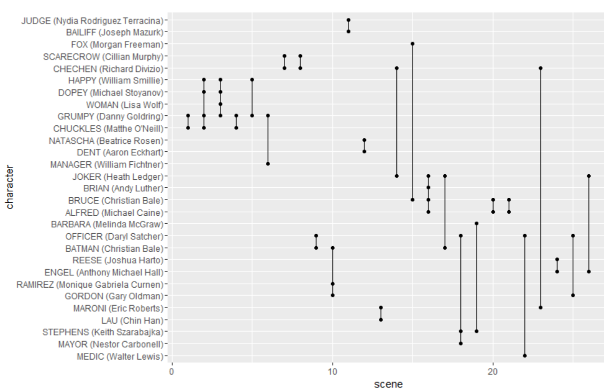


Figure 5: A Line-plot of Characters in *The Dark Knight*

plot to show both the aggregate and dynamic character co-appearances.

Audio

Thirdly, we also visualize the audio of the movie. We hope to find some relationship between the distribution of the sound and the color in the movie. Our process provides a graph of volume distribution over the frames of a movie. Due the size and processing speed concerns, we extracted audio data as signed 16 bits PCM (pulse-code Modulation) WAV file from the DVDs using Audacity. The .wav file is composed of audio samples, which represent the volumes with the amplitude ranging from -1 to 1. Each movie has around three hundred to five hundred million samples, in 48kHz. We choose the same 200 points in time from these samples(the same time stamps as the color palette samples) and visualize their volume distribution in a graph.

RESULTS

In this section we will describe the results obtained from implementing the ideas from our methodology section.

Color Palette

The color palette information was split into genres and decades, and plotted on a Hue versus Value(lightness of the color) graph such as in Figure 6. The area of the square is a direct representation of what percentage that color took up on screen, and the color of the square is a direct representation of the palette color.

Those who wish to interact with this visualization can visit our Jupyter notebook page at `http://goo.gl/rOx2Si`.

This data format provides us with multiple comparisons across genres. For example, in Figure 7, we can see that in general, the color palette of action movies are lower in value than comedy. Not only that, we can also notice that there are more frequent uses of blues and greens, as opposed to colors such as yellow which is prominent in comedies.

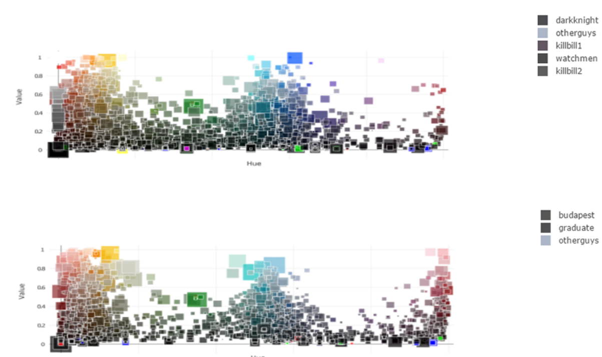


Figure 7: Color Palette plot for Action films versus Comedy films

Another comparison that can be made across decades is shown in Figure 8. You might notice that there is a slight trend towards using more darker blues and brighter yellows in the movies sampled from the 2000s. The colors of 2000s also seem to be more saturated, especially in the reds. This might be accredited to selection bias of more stylized films being selected in the 2000s, or perhaps that there were more saturated and stylized films in the 2000s.

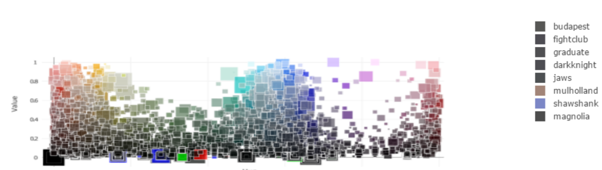


Figure 6: Color Palette plot for Drama films

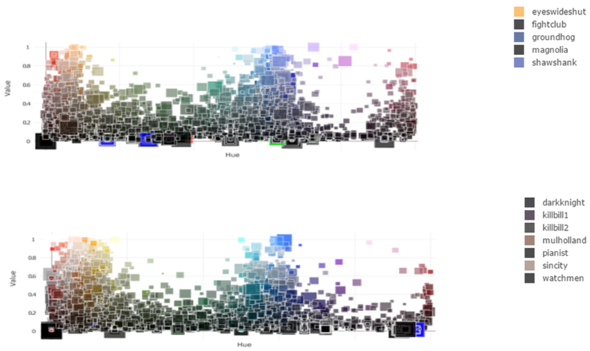


Figure 8: Color Palette plot for 1990s films versus 2000s films

By quantifying the color palette usage and filtering by genre and decades, we hoped to achieve some meaningful way to view trends in movies. As we will discuss further in the Discussion section, we don't make any concrete conclusions from our visualization. We will see the reason why in the next sections as well.

Script Data

Visualizing how characters interact with each other in conjunction with other visual-audio input of the movie can be effective in associating and discovering a certain syntax used in directing of it. For this purpose, we used a network visualization package `networkD3`. The resulting visualization of character co-appearances has leveraged on the idea of David Robinson [15] as shown in Figure 9.

Users can change the scene number using a slider control, and the network visualization and the location of a vertical line in the line-plot changes accordingly. This was also co-located with frame color visualization and audio visualization in such a way to let the users can simultaneously relate the character relations with visual-audio information. The final composition of the tool can be found at https://goo.gl/UTa64R

Audio

To analyze the results of the audio, we divided our collection of nine movies into 3 eras: 1950s–1960s, 1970s–1980s, and 2000s–2010s.

Since the audio are stereo, there are two graphs per movie, each representing the left and right channels of the audio. Our method starts with choosing the first point from the position of the 100th frame of the movie, then the rest of the points are picked with the same increment (the rest of the pixel points over 200). The calculation of picking pixel and drawing of the graphs of volume distribution over 200 frames was processed with SciPy.

From Figure 10 (1950s–1960s), we see that the amplitude of all three movies (Rear Window 1954, The Vertigo 1958, the Graduate 1967) are mostly below 0.1. Most of the value of the points were uniformly distributed in the graph since the range of all the points are between -0.15 to 0.15. Though it could be seen that the volume of the movie stayed constant, there were

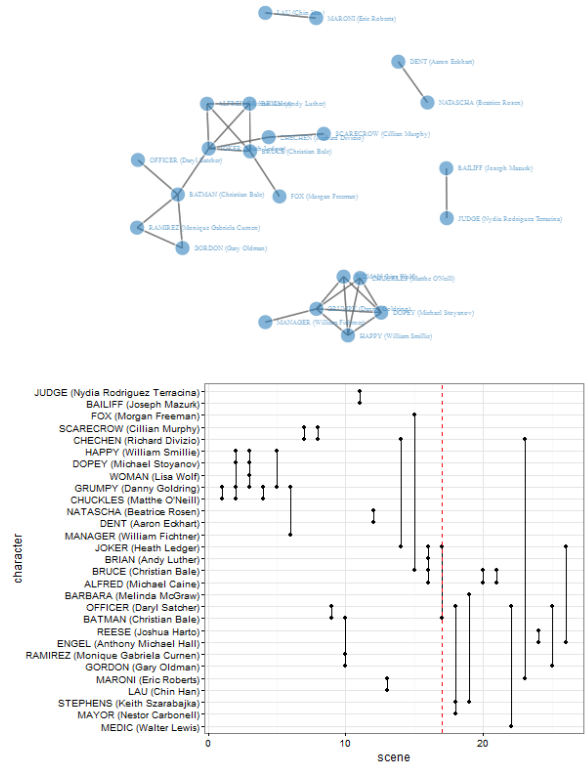


Figure 9: Character Co-appearance Network Visualization over time (*The Dark Knight*)

one or two sharp peaks—indicating a larger amplitude—near the end. We believe this is due to the fact that filmmakers prefer to position the climax in the latter part of the movie. The climax needs the strength of the sound to support its story beat; for example, a fighting scene needs the sounds of guns and fists (Rear Window), or a mystery reveal needs emotional music to entice the audience (Vertigo).

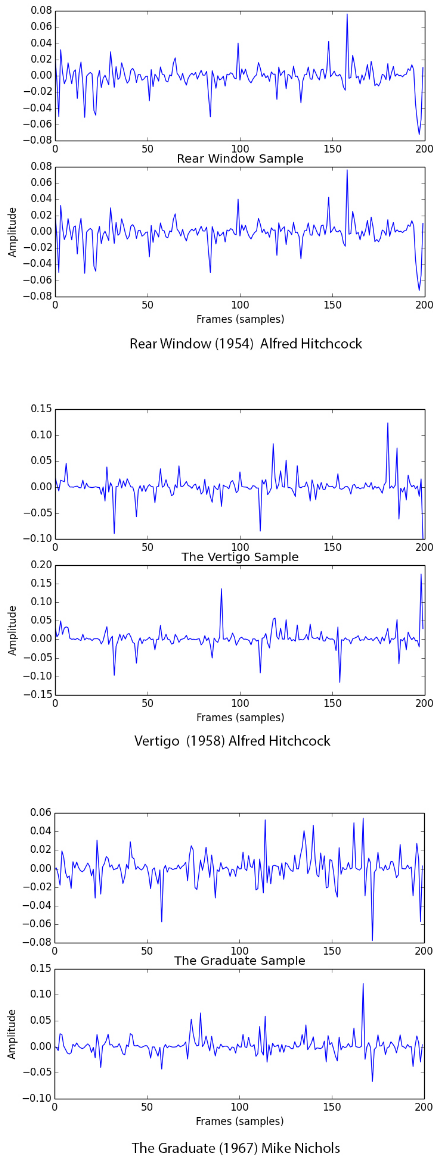


Figure 10: Sound data for movies in the period of 50s and 60s

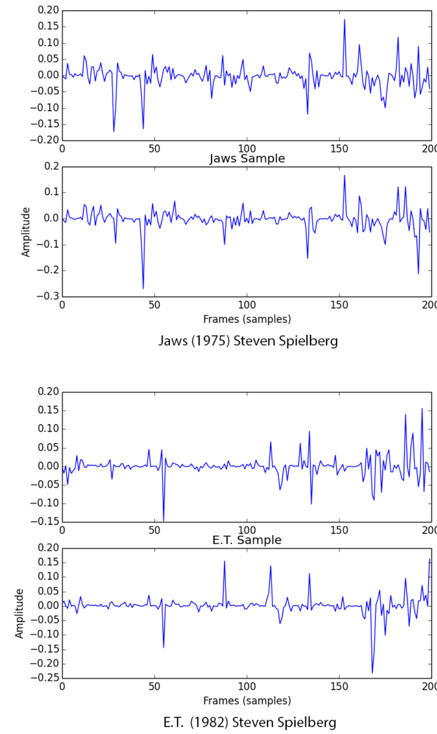


Figure 11: Sound data for movies in the period of 70s and 80s

In Figure 11 (1970s and 1980s), the range of the movies were close to -0.3 to 0.25, which means that the volume of the movies became slightly louder. Also some climactic moments in terms of sound appeared in the former quarter of the movie besides the one showed in the latter parts. In the movie Jaws (1975), the director gives the audience a shocking scene of death with the famous music (the "Jaws theme") in the first twenty minutes, which successfully draws in the attention of the audience, and establishes the theme as a sign of approaching danger.

The audio distribution of the movies from the third periods starts to become more diverse (Figure 12). In the movie Mulholland Drive (2001), the sound has only one climax in the ending part, which accompanies the insanity and suicide of the leading female. With two hours of smooth narration of a dream of the leading character, the sudden climax of the music with a dramatic, tragic ending gave the audience a deep impression. In the other three movies, Kill Bill 1, Kill Bill 2, and the Dark Knight, many peaks of the sound can be found in the graph. Due to all the three being action movies, there were many violent scenes with intense music.

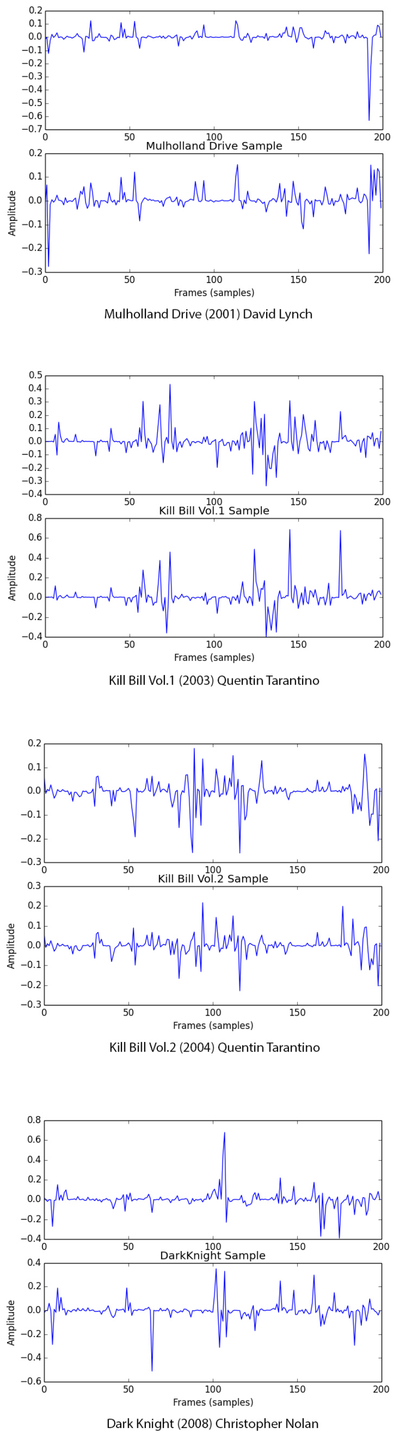


Figure 12: Sound data for movies in the period of 2000s and 2010s

Overall, from all the graphs from these three eras, we can see the tendency of movies becoming louder as they approach the climax, as well as the volume of movies becoming louder in general.

During the process of transforming the audio, we also met several challenges. First of all, we started to transform the audio to be 32 bits PCM format, but the resulting files became so large that we had difficulty processing these files. Also, due to the limitations of time and resources, we could only analyze the audio of nine movies. As a result, our findings in the trend of the audio in the movies might be limited.

DISCUSSION

We set out to fill a gap left by Film Studies in quantifying and visualizing movie data to better study and analyze them in a scientific way. To this end, we looked at three main aspects of film: the color palette, character relations, and audio volume. Though producing interesting results, there are many flaws to the methodology we took.

Firstly, the selection of films were heavily biased by their accessibility and our own taste. This has probably introduced biases such as visually stylistic films being over-represented. A more scientific approach would have sampled some uniform criteria such as release year, critical response to gain a wholesome sample. Another approach could have been to select films of a single criteria such as each year's Academy Award Best Film nominees, that can at least draw a more measured criticism about the selection bias. Instead, our sampling technique was haphazard and requires much improvement in the future. Further, the sheer quantity of our sample films is also insufficient to lead to any conclusions. This is why the conclusions we draw in the Results section are made with caution and disclaimers.

Secondly, our goal of comparing many films resulted in mixed results as we noticed that many films have similar data. For example, in color palettes many films at least have a prominent representation of skin tone. Though skin tone varies heavily depending on the lighting and the LUT applied film to film, they don't stray far from the Red-Orange hues. To differentiate the color palettes, our plan included a transformation on the colors to find which color is more significant in the context of all the films we've gathered, in a manner similar to tf-idf for text. However the math and color perception theory required for this transformation ultimately stopped us from implementing this, and we will revisit this in our next iteration of this work.

Thirdly, despite our criticism of previous works on focusing too much on single movies, our work also doesn't provide a definite answer to this open problem. We wish to implore that further creative research is done in this field to come up with novel visualizations for comparing multiple film data, not simply in small multiples, but a truly easy to view form.

Finally, even though we believe the way we've compiled our data is truly novel, it still doesn't make the narrative structure apparent simply by looking at the data, especially with the more abstract data such as color palette. We have conviction that there must be a way to combine the three different types of data to make a summary of each film that actually makes sense, and will continue to pursue this research to find a visualization that truly makes it easy to consume and review a film for the filmmaker, critic, historian and public.

CONCLUSION

Despite the growth in film—both in consumption and production—the science of studying these films is still lagging behind. As Redfern criticizes, current film studies is performative rather than exploratory, focusing on recitation rather than solving key problems [14]. As more and more film content is produced than ever before, a solid foundation for studying film en masse is required.

By exploring several datasets regarding film, and how they might be visualized for easy consumption, we set out to quantify film data more clearly and succinctly. We hope media researchers will follow suit and that the future of film research helps everyone understand films in a new light.

ACKNOWLEDGMENTS

We thank Professor Jim Hollan for the amazing class COGS220: Information Visualization at University of California, San Diego and our peers for their feedback and support.

REFERENCES

1. 2016. The Internet Movie Script Database (IMSDb). (2016). http://www.imsdb.com/
2. 2016. moviebarcode. (2016). http://moviebarcode.tumblr.com/
3. Frederic Brodbeck. 2016. Cinemetrics. (2016). http://cinemetrics.fredericbrodbeck.de/
4. Liang-Hua Chen, Yu-Chun Lai, and Hong-Yuan Mark Liao. 2008. Movie scene segmentation using background information. *Pattern Recognition* 41, 3 (2008), 1056–1065.
5. Charlie Clark. 2016. The Colors of Motion. (2016). http://thecolorsofmotion.com/films
6. Timothee Cour, Chris Jordan, Eleni Miltsakaki, and Ben Taskar. 2008. Movie/script: Alignment and parsing of video and text transcription. In *Computer Vision–ECCV 2008*. Springer, 158–171.
7. James E Cutting, Kaitlin L Brunick, Jordan E DeLong, Catalina Iricinschi, and Ayse Candan. 2011. Quicker, faster, darker: Changes in Hollywood film over 75 years. *i-Perception* 2, 6 (2011), 569–576.
8. Kevin Ferguson. 2015. Summed frames of Disney films plotted by Median Hue and Median Brightness. (2015). https://www.flickr.com/photos/a2050/22550182823/in/album-72157661482387425/
9. Paul Heckbert. 1982. *Color image quantization for frame buffer display*. Vol. 16. ACM.
10. Go Irie, Kota Hidaka, Takashi Satou, Akira Kojima, Toshihiko Yamasaki, and Kiyoharu Aizawa. 2009. Latent topic driving model for movie affective scene classification. In *Proceedings of the 17th ACM international conference on Multimedia*. ACM, 565–568.
11. Jyotiska. 2014. Extract dominant colors from an image with ColorWeave. (2014). http://blog.dataweave.in/post/83478334651/extract-dominant-colors-from-an-image-with
12. PwC & Website (key4biz.it). 2016. Filmed entertainment revenue in the United States from 2010 to 2019 (in billion U.S. dollars)*. (2016). http://www.statista.com/statistics/259984/filmed-entertainment-revenue-in-the-us/
13. Chao Liang, Yifan Zhang, Jian Cheng, Changsheng Xu, and Hanqing Lu. 2009. A novel role-based movie scene segmentation method. In *Advances in Multimedia Information Processing-PCM 2009*. Springer, 917–922.
14. Nick Redfern. 2014. Quantitative methods and the study of film. Citeseer.
15. David Robinson. 2015. Analyzing networks of characters in 'Love Actually'. (2015). http://varianceexplained.org/r/love-actually-network/
16. Jitao Sang and Changsheng Xu. 2010. Character-based movie summarization. In *Proceedings of the international conference on Multimedia*. ACM, 855–858.
17. Chia-Hung Yeh, Chih-Hung Kuo, and Rung-Wen Liou. 2009. Movie story intensity representation through audiovisual tempo analysis. *Multimedia Tools and Applications* 44, 2 (2009), 205–228.